Modelling and automated calibration of a general multi-projective camera

Khoramshahi, Ehsan

2018-03

Khoramshahi , E & Honkavaara , E 2018 , ' Modelling and automated calibration of a general multi-projective camera ' , Photogrammetric Record , vol. 33 , no. 161 , pp. 86-112 . https://doi.org/10.1111/phor.122

http://hdl.handle.net/10138/237082
https://doi.org/10.1111/phor.12230

THE PHOTOGRAMMETRIC RECORD

The Photogrammetric Record (2018)
DOI: 10.1111/phor.12230

MODELLING AND AUTOMATED CALIBRATION OF A GENERAL MULTI-PROJECTIVE CAMERA

EHSAN KHORAMSHAHI* (ehsan.khoramshahi@nls.fi)

Finnish Geospatial Research Institute, Masala, Finland and *University of Helsinki, Finland*

EIJA HONKAVAARA (eija.honkavaara@nls.fi)

Finnish Geospatial Research Institute, Masala, Finland

**Corresponding author*

Abstract

Recently, multi-projective cameras (MPCs), often based on frame-mounted multiple cameras with a small baseline and arbitrary overlap, have found a remarkable place in geomatics and vision-based applications. This paper outlines the geometric calibration of a general MPC by presenting a mathematical model that describes its unknown generic geometry. A modified bundle block adjustment is employed to calibrate an industrial-level 360° non-metric camera. The structure of any MPC can be retrieved as a calibration set of relative and interior orientation parameters (as well as the pose of the MPC shots) using a calibration room which has been accurately determined by close range photogrammetry. To demonstrate the efficiency and precision of the model, a Panono camera (an MPC with 36 individual cameras) was calibrated. After the adjustment, sub-pixel image residuals and acceptable object-space errors were observed.

KEYWORDS: automatic calibration, bundle block adjustment, multi-projective camera calibration, photogrammetry

INTRODUCTION

WHEN WISHING TO VIEW MORE of the surrounding environment, a multi-camera system (MCS) offers great potential for geomatics instrumentation, robotics, car navigation, entertainment systems and even space applications (such as spacecraft docking navigation systems). Technology based on multiview geometry is inexpensive, accessible and highly customisable. A fixed-structure MCS consists of a set of cameras mounted on a solid platform and is generally simple and flexible enough to focus on the regions of interest. Today, MCSs exist using a wide variety of platforms, some as simple and cheap as 360° cameras and small robots, others being expensive autonomous cars or complex satellite structures.

Recently published literature on MCSs highlights a wide range of real-world and futuristic applications. For example, robotic navigation by stereo vision was proposed by

many researchers such as Grosso and Tistarelli (1993), Desouza and Kak (2002) or English et al. (2014); a multi-camera-based indoor navigation system was successfully demonstrated by Svoboda et al. (2002); Se et al. (2007) proposed a stereo-camera-based 3D modelling of space structures; a multi-camera-based outdoor simultaneous localisation and mapping (SLAM) was demonstrated by Schleicher et al. (2009); a stereo-camera tracking algorithm was successfully tested by Gasparini and Bertolino (2013); multi-camera-based navigation for autonomous cars was discussed and demonstrated by Paracchini et al. (2016); and submarine visual tracking by a stereo camera was shown by Pfingsthorn et al. (2016).

The MCSs that widely appear in recent literature can be categorised according to the architecture of the camera system; for example, categorisation based on the type of lenses incorporated in an MCS:

(1) multi-omnidirectional cameras (MOCs), where the field of view is a full 360°;
(2) multi-projective cameras (MPCs), where the component cameras are more conventional, having a restricted field of view; and
(3) hybrid multiple cameras, with mixed lens mechanisms (for example, including both projective and omnidirectional lenses).

Note that, in this paper, conventional imaging systems which have a restricted field of view are termed *projective* cameras.

It is very logical to have models that either work solely on one category or on more generic models. Many combinatorial cameras, employing different capturing domains, fit into the MPC category, such as multispectral projective cameras or thermal projective cameras. The first two categories simplify the foundation of a calibration process, because both categories only include combinations of the same type of camera, which therefore share a significant degree of calibration modelling, such as interior orientation parameters (IOPs). In contrast, the third category includes different combinations of sensors or camera types on a common platform for image capture. Since an MPC places a limit on the type of lens, a suitable model consequently describes all similar MPCs. In this work, the focus is only on the *calibration modelling of a general MPC.*

Single-frame Cameras and Stereo Cameras $MCS_n(n=2)$

For many decades, conventional single-frame cameras (SFCs) and fixed-base stereo cameras (two cameras, often mounted on a bar with a known separation) (MPC_2 or MOC_2) have been successfully employed in both close range photogrammetry and computer vision applications, mainly because of their design simplicity and ease of calibration. Many diverse applications of stereo cameras have been proposed in the literature. Recent applications of SFCs and stereo cameras may be classified into three main categories:

(1) 3D structure-and-motion estimation, for example, dense reconstruction (Furukawa and Ponce, 2007; Furukawa et al., 2010), sparse bundle adjustment and Internet-based reconstruction (Snavely et al., 2008), space structure estimation (Se et al., 2007), incremental structure from motion (Wu, 2013), stereo panoramas (Amini et al., 2014) and dual fluoroscopy imaging system calibration used for bone reconstruction (Lichti et al., 2015).
(2) Navigation systems, for example, indoor localisation (Svoboda et al., 2002), stereo tracking for dental surgery (Wang et al., 2014), car trajectory estimation (Paracchini et al., 2016), underwater navigation (Pfingsthorn et al., 2016) and spacecraft precise docking (Tweddle, 2010).

(3) Combinatory applications with other sensors, for example, outdoor SLAM (Schleicher et al., 2009).

In most above-mentioned applications (such as indoor positioning), MCSs offer the unique possibility of rapidly capturing 360° shots. The variety of commercially available sensors makes it feasible to measure several spectral bands of data as SFC output. A rigid body usually ensures that the relative locations of cameras (with respect to a local coordinate system) are fixed, at least during a measurement period. A brief look at the development of recent imaging systems reveals that MCSs are rapidly growing to become a leading imaging technology in many diverse areas. In order to achieve the best geometric use of a camera, the applications upon which the system is founded need to be carefully considered. However, because of the shared properties of MPCs, a universal paradigm could be defined for the calibration phase.

Multi-camera Systems ($MCS_n, n > 2$)

Multiview systems, where $MCS_n(n > 2)$, are usually developed for applications for which two-image stereo vision is not flexible enough to capture the environment surrounding a moving platform. A few examples of such cases are 360° multi-camera panoramic imaging sensors, vehicle autonomous navigation cameras, indoor cameras and multiview reconstruction.

Based on applications that an $MCS_n(n > 2)$ is designed for, cameras are mounted in a wide variety of geometric configurations; therefore, a system calibration method needs to be flexible enough to work with any configuration. In recent literature, two main configurations have been considered for a multiview $MCS_n(n > 2)$:

(1) The cameras are mounted solidly on an object or a piece of apparatus (such as the walls of a building or a multi-camera rig).
(2) The cameras are mounted on a solid, but movable, frame (for example, a multi-camera semi-panoramic system).

Since the topology and dynamic properties are different between these two groups, different calibration paradigms need to be defined for each group. In this work, the focus is on MPCs of the second group, where a movable MPC consists of a set of cameras fixed on a frame (usually pointing outwards).

State-of-the-art of MPC Calibration

A standard bundle block adjustment (BBA) has been proven to be an efficient tool for metric single/stereo camera calibration (for example, Granshaw, 1980; Fraser, 1997, 2013; Gruen and Huang, 2001). A photogrammetric BBA was then used to calibrate non-metric cameras in a very cost-effective manner, for example, Triggs (1998) proposed an automatic calibration approach coupled with planar coded targets for projective cameras. His method was able to run the calibration with five or more image exposures. His only assumption was that the IOPs and the structure remained fixed during measurement. Zhang (2000) introduced a planar chequerboard coded target to calibrate a projective camera, employing a corner detector to automatically find the coded target in an image set. A few images from different angles were usually enough to run the BBA. Bouguet (2011) later modified Zhang's toolbox to accept stereo cameras. Both these methods have demonstrated deficiencies when it comes to comprehensive, scene-independent photogrammetric (as

opposed to computer vision) calibration. Camera calibration has been undertaken for many decades; Clarke and Fryer (1998) provide a useful historical overview from a photogrammetric standpoint. Some very early attempts to calibrate a stereo camera from a computer vision perspective appeared in works such as those of Zhuang (1995), who used distance measurements of a fixed-length moving object such as a utility pole to find the extrinsic (external orientation) parameters of a stereo camera. Using a different methodology, Lerma et al. (2010) employed distance constraints between camera perspective centres to calibrate an MPC_3 consisting of a stereo camera coupled with a thermal camera, using a rigid MPC body as a part of a multi-sensor system. In their work, high accuracies were achieved only by employing a full set of baseline distance constraints. Svoboda et al. (2002) used an easily detectable moving object (a laser track pointer) to calibrate a set of cameras that were fixed in an indoor environment. Their goal was to find interior and exterior orientation parameters (IOPs and EOPs) of a set of cameras that was specifically designed for the indoor localisation of moving objects. In the MPC_n of Svoboda et al. (2002), the distances between cameras were relatively large, and the MPC_n was supposed to stay geometrically fixed during the operational period of the localisation system (see the toolbox of Svoboda et al., 2011). Tommaselli et al. (2014) designed a terrestrial calibration field where 139 AURUCO coded targets (Garrido-Jurado et al., 2014) were installed on the floor and side walls of a room, modifying the collinearity equations to accept their proposed fisheye lens model.

The variation of IOPs and EOPs of a stereo camera and an MPC was reviewed by Habib et al. (2014). Modified collinearity equations for an MPC by enforcing the relative pose (position and orientation) of all cameras with respect to a reference camera/frame have been demonstrated, for example, in works of He et al. (1993), Tommaselli et al. (2013) and Habib et al. (2014). This formulation is listed in Habib et al. (2014) as a "*one step procedure*" where invariant relative orientation parameter (ROP) enforcement is applied at different times, and it "*intensifies with the increase of the number of cameras*"; however, they used the modified collinearity formulation for stability analysis of an experiment with an MPC. In their work, mean and standard deviation values of time-dependent calibration parameters were directly calculated from time series.

Li et al. (2013) proposed an algorithm that was able to detect a partially visible chequerboard coded target. Consequently, their method was usable even for cases when the field of view of neighbouring cameras did not overlap. They mainly focused on projective and catadioptric cameras (that use a combination of lenses and curved mirrors) MPC_n, and used geometrically known coded targets to connect neighbouring cameras, under the condition that no overlap exists. In their paper, an MPC_4 of projective cameras on a rig with a separation angle of 90° was calibrated. Urban et al. (2017) presented a BBA approach (called MultiCol) to calibrate an MOC_n. They simulated an MOC_{15} with a few randomly placed cameras in order to show the usefulness of their method and implementation.

The lack of a uniform photogrammetric paradigm for a multiview $MPC_n (n \geq 2)$ is understandable. Diverse technical language has been used in recent literature. The most significant literature on this topic could not be seen as sufficient and general enough for the MPC_n studied in this paper (Svoboda et al., 2002; Li et al., 2013; Habib et al., 2014; Urban et al., 2017). In some cases, the focus is limited to the special condition that a specific MPC/MOC is designed for (Svoboda et al., 2002). This current work, however, tries to continue the aforementioned work in terms of photogrammetric formulation integrity, as well as the practicalities that are involved in a real multiview calibration process with an $MPC_n (n \geq 2)$. In this paper a modified BBA approach is demonstrated for calibrating a general MPC, based on the customised collinearity equations defined by He et al. (1993),

Tommaselli et al. (2013) and Habib et al. (2014). The proposed formulation is employed to improve the methodology stated by Li et al. (2013) and Urban et al. (2017). The main improvements over this previous work are: (i) ease of calibration through automation; (ii) error propagation based on a standard statistical model; and (iii) formulating the problem as a standard photogrammetric paradigm. This work demonstrates an efficient, practical and precise methodology based on non-linear least squares estimation for calibrating a general MPC.

The method presented in this work is based on the following six factors:

(1) Taking a set of images from a calibration room filled with coded targets with an SFC to precisely retrieve the positions of the targets and the corresponding uncertainties by employing a minimum-constraint BBA.
(2) Taking a second set of images of the calibration room with an MPC.
(3) Executing a minimum-constraint BBA for individual cameras of the MPC.
(4) Approximating an initial structure for the MPC and the initial pose (locations and orientations) of images by combining the output from the individual BBA runs.
(5) Enforcing the structure of the MPC in a standard BBA as relative orientation constraints within the collinearity model to build a customised BBA for a general MPC.
(6) Discovering the internal structure of the underlying MPC, IOPs of individual cameras, positions and orientations of the MPC images by running the customised BBA whilst assuming the coded target locations fixed.

The structure of this paper is as follows. Theoretical aspects are described first, followed by specifications of the hardware (such as the calibration room and cameras). The method is presented subsequently. The results are demonstrated and discussed afterward, and finally the paper is summarised. Symbols and conventions that are used in this paper are listed in the Appendix.

Theoretical Aspects of the MPC Calibration

In this section, the required technicalities are expanded in photogrammetric terminology to address the MPC calibration problem. The intention is to keep the formatting standard and consistent with well-known photogrammetric work, such as that of Gruen and Huang (2001). This section starts with basic stereographic formulations and leads to a complete MPC calibration paradigm.

Coplanarity and Collinearity Equations

The coplanarity and collinearity equations are fundamental to photogrammetric adjustment. The coplanarity equation expresses an ideal assumption that, in a single stereopair, an object point, its two corresponding image points in the stereopair and the perspective centres (nodal points) of both camera positions lie on a plane. In contrast, the collinearity equations are defined for an individual image, whereby an object point, the perspective centre and the corresponding image point (on the image plane) lie on the same straight line.

Fig. 1(a) depicts a stereopair of an SFC whereby each image is captured with the same sensor but at a different time. The local coordinate system of an image is demonstrated in Fig. 1(b). In this work, it is assumed that the origin of the local coordinate system of an image is the perspective centre, where the x and y axes are parallel to the image plane, and

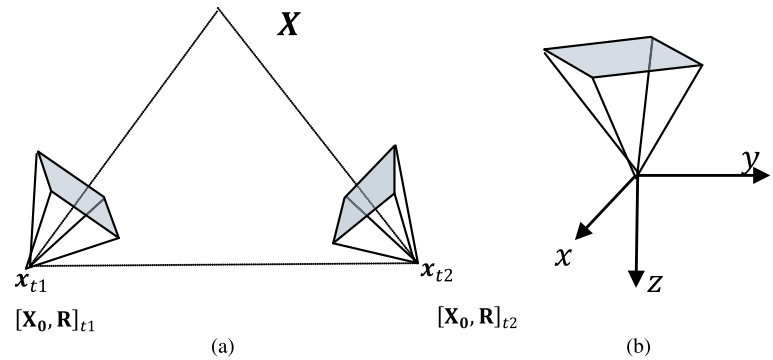


FIG. 1. Coplanarity, collinearity and a camera's local coordinate system. (a) A central perspective (pinhole) camera looking towards to an object point X at epochs $t1$ and $t2$ – the coplanarity equation is defined based on the intersection geometry. (b) The local coordinate system of an ideal (pinhole) camera.

the z axis is perpendicular to the image plane and positive in the direction towards the object space.

Coplanarity Equation. The coplanarity condition, as depicted in Fig. 1(a), can be formulated as a vector product in the following form:

$$(\mathbf{X}_0)_{t1}.((\mathbf{X}_0)_{t2} \times \mathbf{X}) = 0 \tag{1}$$

where $(\mathbf{X_0})_{ti}$ is the position vector of the camera at time (ti). In equation (1), the products are the inner product (.; also termed the dot or scalar product) and the outer product ($\times$; also termed the cross or vector product), respectively. Equation (1) could, equivalently, be rewritten as a matrix product:

$$(\mathbf{X_0})_{t1}^{T}.[(\mathbf{X_0})_{t2}]_x.\mathbf{X} = 0 \tag{2}$$

where $[(\mathbf{X_0})_{t2}]_x$ is a 3×3 matrix of the form:

$$\begin{pmatrix} 0 & -\lfloor(\mathbf{X}_0)_{t2}\rfloor_3 & \lfloor(\mathbf{X}_0)_{t2}\rfloor_2 \\ \lfloor(\mathbf{X}_0)_{t2}\rfloor_3 & 0 & -\lfloor(\mathbf{X}_0)_{t2}\rfloor_1 \\ -\lfloor(\mathbf{X}_0)_{t2}\rfloor_2 & \lfloor(\mathbf{X}_0)_{t2}\rfloor_1 & 0 \end{pmatrix}. \tag{3}$$

For constructing a network of images with unknown locations and orientations, the coplanarity equation can be helpful in situations where the collinearity equations could easily erroneously converge (get stuck) to local optimums if the initial values are not close enough to global optimums.

Collinearity Equations. After finding sufficiently close approximations of the pose (location and orientation) variables, the collinearity equations play an essential role by removing all the unnecessary constraints and connecting all cameras throughout a uniform model. If we assume a simple central perspective (pinhole) camera model with a local coordinate system that looks downward, the collinearity condition is formulated as the following:

$$\mathbf{X} = \Lambda.\mathbf{R}_{(\omega,\phi,\kappa)_t}.\mathbf{x} + (\mathbf{X_0})_t. \tag{4}$$

In equation (4), Λ is an unknown scale factor and $\mathbf{R}_{(\omega,\phi,\kappa)_t}$ is the Euler rotation matrix of a camera at a certain time t. Using the above notion, an *essential matrix* $\mathbf{E}$ is defined as:

$$\mathbf{E} = \left[(\mathbf{X_0})_t\right]_x.\mathbf{R}_t. \tag{5}$$

If the unknown scale (Λ) in equation (4) is removed, observation equations are generated that will be employed in the body of the BBA:

$$\begin{aligned} x_j.\left[\mathbf{M}_{t(3,:)}.(\mathbf{X} - (\mathbf{X}_0)_t)\right] + \mathbf{M}_{t(1,:)}.\left(\mathbf{X}_j - (\mathbf{X}_0)_t\right) = 0, \\ y_j.\left[\mathbf{M}_{t(3,:)}.\left(\mathbf{X} - (\mathbf{X}_0)_t\right)\right] + \mathbf{M}_{t(2,:)}.\left(\mathbf{X}_j - (\mathbf{X}_0)_t\right) = 0 \end{aligned} \tag{6}$$

where $\mathbf{X}_j$ is the corresponding object point for the image point $\mathbf{x}_j$, and $\mathbf{M}_t = \mathbf{R}_t^{-1}$.

In equations (4) to (6), perspective centre (pinhole) coordinates $\mathbf{X}_0$ (Fig. 1) are used. The lens distortion and scale factor inherent in a real optical system are considered as a non-linear function that connects the actual pixel coordinates to corrected image coordinates (Brown, 1966; Fraser, 1982; Gruen and Huang, 2001):

$$\begin{aligned} x_1 = \frac{(x)_{t_1} - PP_x}{f}, y_1 = \frac{(y)_{t_1} - PP_y}{f}, r = \sqrt{x_1^2 + y_1^2}, Rad = \left(1 + K_1.r^2 + K_2.r^4 + K_3.\mathrm{r}^6\right), \\ (x_n)_{t_1} = x_1.Rad + 2.P_1.x_1.y_1 + P_2.\left(r^2 + 2(x_1)^2\right) - \delta.x_1 + \lambda.y_1, \\ (y_n)_{t_1} = y_1.Rad + 2.P_2.x_1.y_1 + P_1.\left(r^2 + 2(y_1)^2\right) + \lambda.x_1 \end{aligned} \tag{7}$$

where $(x_n)_{t_1}, (y_n)_{t_1}$ are the undistorted image coordinates corresponding to the distorted pixel coordinates $((x)_{t_1}, (y)_{t_1})$; *Rad* is the radial term containing the symmetric radial distortion coefficients (K_1, K_2, K_3); P_1 and P_2 are the decentring distortion coefficients; (PP_x, PP_y) and f are the location of the principal point and the principal distance in pixel units, respectively; and δ, λ are scale and shear factors, respectively.

The Normalised 8-point Algorithm

In this work estimating an essential matrix is achieved using the normalised 8-point algorithm (Hartley, 1997). If it assumed that $\mathbf{x}_{t_1}$ is the location of an image point in frame (t_1) in pixel units, this point is transferred to the corresponding corrected perspective centre (pinhole) image coordinates $\bar{\mathbf{x}}_{t_1}$ by using equation (7). For a set of normal image points, normalised "normal image points" (forming *normalised image points*: see Hartley, 1997; Hartley and Zisserman, 2004) are calculated based on the centroid of the points and scaling to one:

$$\bar{\bar{\mathbf{x}}}_{t_1} = s.(\bar{\mathbf{x}}_{t_1} - \bar{\mathbf{M}}_{t_1}) \tag{8}$$

where $\bar{\mathbf{M}}_{t_1}$ is the mean of the normalised image points and S is the scale normalising term. The coplanarity equation is therefore rewritten in the following form:

$$\left(\bar{\bar{\mathbf{x}}}_{t_1}\right)^T.\bar{\mathbf{E}}_{(t_1,t_2)}.\left(\bar{\bar{\mathbf{x}}}_{t_2}\right) = 0. \tag{9}$$

The estimated essential matrix, $\bar{\mathbf{E}}_{(t_1,t_2)}$, encodes the relative orientation connection between these two image frames. This equation is equivalent to a linear system of equations:

$$\mathbf{A}_{(m\times 9)} \,.\, \bar{\mathbf{E}}_{(t_1,t_2)(9\times 1)} = \mathbf{0}. \tag{10}$$

One solution to the above equation is retrieved by finding the singular-value decomposition (SVD) of $\mathbf{A}$. The estimated $\mathbf{E}$ is finally decomposed into a relative rotation matrix $[\mathbf{R}_{t_2} \,.\, \mathbf{R}_{t_1}^{\mathrm{T}}]$ and a positional vector $[(\mathbf{X}_0)_{t_2} - (\mathbf{X}_0)_{t_1}]$.

The Self-calibrating BBA of a Multi-projective Camera

A single image of an MPC$_n$ is represented in this work by the following parameters:

(1) *Inner orientation parameters:*
$\mathbf{IOP}_i \; (i = 1{:}n)(f, PP_x, PP_y, K_1, K_2, K_3, P_1, P_2, \delta, \lambda)$.

(2) *Relative orientation parameters*:
$\mathbf{ROP}_{(i)}(i = 2 : n)\{(\zeta, \eta, \psi), (\Delta_x, \Delta_y, \Delta_z)\}$; the first set of parameters $(\zeta, \eta, \psi)_i$ are the relative Euler angles of the ith sensor with respect to the first sensor in the MPC$_n$, and the shift vector $\left(\Delta_x, \Delta_y, \Delta_z\right)_i$ is the relative position of the ith camera with respect to the first camera in the local coordinate system of the MPC$_n$.

(3) *Pose*:
The location and orientation of the frame at the moment of image capture $(\mathbf{R}_t, (\mathbf{X}_0)_t)$ with respect to the local coordinate system.

Therefore, the number of calibration parameters (*nCalib*) for an MPC$_n$ is:

$$nCalib = n \times 10 + (n-1) \times 6 = 16 \times n - 6. \tag{11}$$

The results of "The Normalised 8-point Algorithm" section provide an initial estimation for the pose (locations and orientations) of a set of cameras with respect to a local object coordinate system. The next step is to find a more stable estimation by applying the least squares approach. The main equations that are employed for this are based on the collinearity of an object point, a perspective centre and the corresponding image point. Since a given object point potentially emanates to image points in multiple photographs, much better estimations are feasible compared to coplanarity estimations. Collinearity in its basic form (equations (4) and (6)) provides a non-linear function of the calibration parameters, object points and corresponding image points. In an MPC$_n$, the observation equation (6) can be rewritten as:

$$\begin{aligned} F_{x_j} &= x_j.\left(\mathbf{M}_{i,t}\right)_{(3,:)}.\left[\mathbf{X} - \Delta_{i,t}\right] + \left(\mathbf{M}_{i,t}\right)_{(1,:)}.\left[\mathbf{X} - \Delta_{i,t}\right] = 0, \\ F_{y_j} &= y_j.\left(\mathbf{M}_{i,t}\right)_{(3,:)}.\left[\mathbf{X} - \Delta_{i,t}\right] + \left(\mathbf{M}_{i,t}\right)_{(2,:)}.\left[\mathbf{X} - \Delta_{i,t}\right] = 0 \end{aligned} \tag{12}$$

where $\mathbf{M}_{i,t} = \mathbf{R}_{\mathbf{i,t}}{}^{-1}$ and $\mathbf{R}_{i,t}$ is the rotation matrix of the ith camera at time t when the orientation of the frame is $(\omega, \phi, \kappa)_t$. The calculation of $\mathbf{R}_{i,t}$ uses:

$$\mathbf{R}_{i,t} = \mathbf{R}_{(\omega,\phi,\kappa)_t}.\mathbf{R}_{(\zeta,\eta,\psi)_i}(i = 2 : n). \tag{13}$$

In a similar way $\boldsymbol{\Delta}_{i,t}$ is defined as:

$$\boldsymbol{\Delta}_{i,t} = \mathbf{R}_{(\omega,\phi,\kappa)_t}.\boldsymbol{\Delta}_i + (\mathbf{X}_0)_t \tag{14}$$

where $(\mathbf{X}_0)_t$ is the location of the MPC_n at time t in object space. Subsequently, equation (14) boils down to a non-linear function of the form:

$$\mathbf{F}\Big((f, \mathbf{PP}, \mathbf{K}, \mathbf{P}, \delta, \lambda)_{i=1:n}, \mathbf{R}_{(\zeta,\eta,\psi)_{i=2:n}}, \boldsymbol{\Delta}_{i=2:n}, R_{(\omega,\phi,\kappa)_{t=t_1:t_m}}, (\mathbf{X}_0)_{t=t_1:t_m}, \mathbf{X}_{(1:n\mathrm{O})}, \mathbf{x}_{(1:n\mathrm{I})}\Big) = 0. \tag{15}$$

The modified collinearity equations (12) enforce the internal structure of an MPC into collinearity equations representing the relative position and orientation of projective cameras with respect to the local coordinate system of the first camera in the MPC. Subsequently, equation (15) combines all non-linear observation equations to be the basis for the least squares estimation. As a result, the output reflects the structure of the MPC in addition to the inner orientation parameters, and positions and orientations of the camera stations.

Bundle Block Adjustment

In this work, the statistical model used for non-linear adjustment is based on the combined adjustment method that was discussed by Wells and Krakiwsky (1971). Primarily a set of non-linear functions is considered:

$$\mathbf{f}(\mathbf{x}, \mathbf{l}) = \mathbf{0} \tag{16}$$

where $\mathbf{f}$ is m non-linear functions of the observations $\mathbf{l}$ and unknowns $\mathbf{x}$. A first-order Taylor expansion of $\mathbf{f}$ around an initial approximation is then expressed as:

$$\mathbf{f}(\mathbf{x}, \mathbf{l}) = \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}\mathbf{x}}\mathrm{d}\mathbf{x} + \frac{\mathrm{d}\mathbf{f}}{\mathrm{d}\mathbf{I}}\mathrm{d}\mathbf{I} + \mathbf{f}(\mathbf{x}_0, \mathbf{l}_0) = \mathbf{0} \tag{17}$$

where $\frac{\mathrm{d}\mathbf{f}}{\mathrm{d}\mathbf{x}}$ is the partial derivative of $\mathbf{f}$ with respect to $\mathbf{x}$ (Jacobian matrix with respect to the unknowns) and $(\mathbf{x}_0, \mathbf{l}_{l0})$ are the initial estimations of the unknowns and observations. After linearisation, the mathematical model of the combined method comprises of m observation equations that can be expressed as:

$$\mathbf{A}.\mathbf{dx} + \mathbf{B}.\mathbf{v} + \mathbf{w} = \mathbf{0} \tag{18}$$

where $\mathbf{A}$ is the Jacobian matrix with respect to the unknowns, $\mathbf{B}$ is the Jacobian matrix with respect to the observations, $\mathbf{v}$ is the residual vector of observations and $\mathbf{w}$ is the misclosure vector:

$$\mathbf{w} = \mathbf{f}(\mathbf{x}_0, \mathbf{l}_0). \tag{19}$$

Either an analytical or a numerical estimation of the Jacobian matrix could be fed into equation (18); here, the latter form is formulated and employed. The optimisation step is formulated as:

$$\begin{bmatrix} J_{11} & \ldots 0 & J_{12} \ldots & 0 & J_{13} \ldots & 0 & J_{14} & \ldots & 0 \\ \vdots & & & & & & & & \vdots \\ 0 & \ldots J_{m1} & 0 & \ldots & J_{m2} & 0 & \ldots & J_{m3} 0 & \ldots & J_{m4} \end{bmatrix} \begin{bmatrix} \mathrm{d}.\mathbf{IOP}_1 \\ \vdots \\ \mathrm{d}.\mathbf{IOP}_n \\ \mathrm{d}.\mathbf{Rel.EOP}_1 \\ \vdots \\ \mathrm{d}.\mathbf{Rel.EOP}_n \\ \mathrm{d}.(\omega,\phi,\kappa)_1 \\ \vdots \\ \mathrm{d}.(\omega,\phi,\kappa)_t \\ \mathrm{d}.(\mathbf{x}_0)_1 \\ \vdots \\ \mathrm{d}.(\mathbf{x}_0)_t \end{bmatrix} + \mathbf{B}.\mathbf{v} + \mathbf{w} = 0. \quad (20)$$

Calibration parameters

EOPs

A variation function is defined, based on Lagrange coefficients, to carry out a least squares estimation. In the above system of linearised equations, the least squares estimation of the unknowns is a Newton step toward the location of a local extremum:

$$[\mathrm{d}\hat{\mathbf{x}}] = [\mathbf{A}^{\mathrm{T}}.(\mathbf{B}.\mathbf{P}^{+}.\mathbf{B}^{\mathrm{T}})^{+}.\mathbf{A}]^{-1}.[\mathbf{A}^{\mathrm{T}}.(\mathbf{B}.\mathbf{P}^{+}.\mathbf{B}^{\mathrm{T}})^{+}.\mathbf{w}] \quad (21)$$

where **P** is a positive definite matrix of observation weights. Sparsity in the Jacobian matrices **A**, **B** and weight matrix **P** of equations (18), (20) and (21) are employed to facilitate the weighted non-linear least squares solution, otherwise a considerable amount of memory is required to solve the adjustment.

MATERIALS

In this section, different aspects of the experiment's materials, such as specifications of the calibration room, the SFCs and the MPC, are given.

Single-frame Cameras

In order to determine the geometry of the calibration room, an SFC was employed (Fig. 2(a)). The camera used was a Canon EOS 6D digital single-lens reflex (DSLR) camera, with a 20 Mpixel (5472 × 3648) complementary metal oxide semiconductor (CMOS) detector, a Canon EF 24 mm f/2·8 IS USM lens with a principal distance (calibrated focal length) of 20·650 mm ± 2 μm and a field of view (FOV) of $81{\cdot}82° \pm °1'$ (Fig. 2(a)). A second SFC (Fig. 2(b)) was employed for cross-checking the calibration room's coded target locations. This was a Samsung NX300 mirrorless digital camera, again with a 20 Mpixel (5472 × 3648) CMOS detector, a Samsung ultra-wide-angle f/2·4 lens with a principal distance of 16·340 mm ± 3·3 μm and a FOV of $71{\cdot}41° \pm °1'$ (Fig. 2(b)). Approximate initial estimations of the IOs were used to initiate the BBA.

MPC Camera

A Panono camera, a commercially available MPC manufactured by Professional360 (Berlin, Germany), was chosen to show the capabilities of the underlying approach. A

(a) (b)

FIG. 2. Single-frame cameras: (a) Canon EOS 6D; (b) Samsung NX300.

Panono is a throwable multi-projective semi-panoramic camera ball with 36 projective cameras that look outward (Fig. 3). The camera was introduced in 2014 as a successful start-up project.

After stitching using cloud software, the Panono provides stunning, non-metric, 108 Mpixel, 360° images. The output of the cloud software is mainly suitable for non-metric applications, such as panoramic visualisation. Each neighbouring pair of cameras has a very small (<10%) overlap. The MPC has three main structural surfaces that lock together to form a camera ball. Each surface contains 12 projective cameras, mounted on a plastic frame. Each projective camera consists of a red/green/blue (RGB) sensor with 2064×1552 pixels. The approximate focal length of each projective camera is the equivalent of about 1900 ± 30 pixels, the FOV is about $57° \pm 0{\cdot}4°$ and pixel size is $3{\cdot}07\,\mu m$. A Panono (an Explorer Edition [2015]) was used for the MPC calibration.

For the calibration, a room measuring $356\,cm \times 519\,cm \times 189\,cm$ was considered as the rigid structure. A total of 215 coded targets were printed and attached to the ceiling, floor, walls and a staircase railing. Since the attached coded targets were not strictly rigid, additional laminated dots were printed and fixed to ensure the rigidity of the targets (Fig. 4).

Image Datasets

Seven datasets, each containing between 50 and 90 images, were captured by the Canon EOS 6D camera with a ground sampling distance (GSD) of between 0·41 and 1·03 mm. Two cross-check datasets of 50 and 80 images were taken by the Samsung NX300 camera (GSD between 0·20 and 0·62 mm).

Calibration Room

For the Panono's calibration, two datasets were captured and analysed. The first of these contained 34 panoramic exposures, and thus 1224 (34×36) individual images at three different height levels (adjusted by a tripod); the GSD was between 0·32 and 0·86 mm. The second dataset contained 84 panoramic shots, providing 3024 (84×36) individual images at two different height levels; the GSD in this case was between 0·20 and 0·67 mm. At the

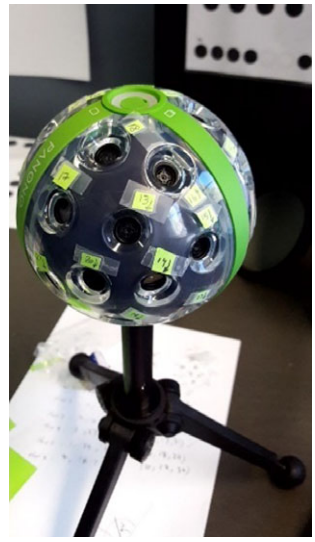
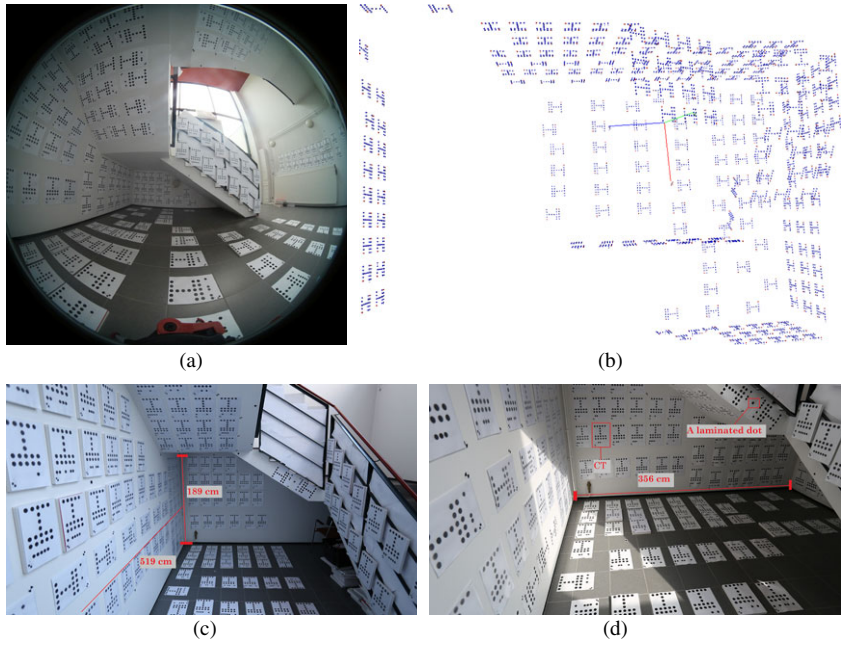


FIG. 3. The Panono camera; all 36 projective cameras are individually labelled.



(a) (b) (c) (d)

FIG. 4. (a) A 360° view of the calibration room at the Finnish Geospatial Research Institute (FGI) with the installed targets. (b) The full set of 3D targets. (c) Side view and dimensions of the calibration room. (d) Coded targets (CTs) and robust laminated dots installed on all surfaces.

lowest height level, the camera was fixed on a line slider in such a way that every 5 to 10 shots were taken in a straight line. This physical constraint was later found useful in checking the adjusted motion.

Method

In this section, different aspects of the methodology are discussed, including initialisation steps together with hardware and software implementation.

MPC Bundle Block Adjustment Design

Firstly, by assuming the room to be a fixed structure, a standard BBA is performed for every individual projective camera of the Panono. The output of the BBA provided approximate motions of cameras with respect to the local coordinate system. Implicit scale bars were applied by employing fixed distances to the internal structure of the coded targets. Since the extracted motion of the projective cameras could be fairly noisy in the case of the Panono (with its non-metric projective cameras), weighting was applied in order to reduce the effect of noise. The estimated IOPs and EOPs of individual cameras were also a side product of running an individual BBA. By combining the resultant "projects" (containing 36 individual project files), a uniform project containing all 36 component cameras was formed. The next step was to estimate the initial structure of the Panono (MPC_{36}) by combining the resultant relative poses (orientation and position) of the cameras with respect to the reference projective camera. The initial structure was estimated by calculating the geometric mean of individual BBA outputs to reduce the effect of noise as much as possible:

$$(\zeta_0, \eta_0, \psi_0)_i = \frac{1}{m} \sum\nolimits_{s=1:s_m} (\zeta, \eta, \psi)_s \tag{22}$$

where m is the number of cameras and s represents the individual camera.

The resulting MPC_{36} sensor resembled the initial structure of the underlying sensor. Finally, the MPC_{36} structure was retrieved by optimising the least squares cost function by employing Newton's iterative solution (see the section "The Self-calibrating BBA of a Multi-projective Camera"). The result of this process is a least square estimate of the underlying calibration parameters, their corresponding variance/covariance values and image residuals of the observations. These results depicted the structure of the underlying sensor.

Software and Hardware

Equation (15) was linearised and coded in C++. A 64-bit Microsoft C++ compiler was employed to build executable binaries with the possibility of accessing large portions of random-access memory (RAM). A computer with a Core i7 4712HQ processor with 16 GB of RAM was used for data processing. Qt (4.8), Coin3D (3.1.3) and SoQt (1.5) were employed for the graphical user interface (GUI) and 3D visualisations. An optimised implementation of LAPACK (Intel Parallel Studio IPP 2015) was used for the matrix operations such as multiplication and factorisation (SVD and QR decomposition). Whenever possible, multi-threading was employed within the code by adopting OpenMP tags.

Automatic Tie-point Extraction Approach

The customised coded targets employed in this work are based on Khoramshahi et al. (2017). The structural properties of the customised coded targets have enabled the implementation of the automatic detection algorithm in MATLAB. The main pipeline of this algorithm consists of:

(1) Running the maximally stable extremal region (MSER) algorithm to find approximate locations of ellipses (that approximate MSER regions).
(2) Applying *k*-means clustering on the extracted region to find boundaries with sub-pixel accuracy.
(3) Fitting rotated ellipses to extracted boundaries.
(4) Clustering extracted ellipses.
(5) Using the topological structure of the coded target to find targets in clusters of ellipses.
(6) Establishing a projective transformation between ideal 2D locations of coded targets and their corresponding location in images to automatically read the identification (ID) part.

Performance Assessment

The accuracy of the 3D-point reconstruction using the Panono camera was assessed using two different approaches. In the first approach, 2300 reference 3D check points were selected from Canon EOS 6D dataset 7. The standard deviations of the points from this dataset's adjustment were observed to be less than 0·7 mm in all directions. The reference labels were excluded from the Panono adjustment; therefore, the result of this adjustment were independent of these points. By employing the retrieved EOP/IOP parameters of the Panono's adjustment, a corresponding set of 3D coordinates for these labels were estimated by spatial intersection of image points. Two sets of coordinates (from the Canon EOS 6D and the Panono) were then co-registered and the residual values were used as the first performance metric. In the second approach, 150 scale bars from intersected object points in Panono dataset 2 were compared with a priori known values. Since the coded-target structure was known, the longest distance could be calculated from the ideal size of an A4 piece of paper and used as the reference value for the scale bars.

RESULTS AND DISCUSSION

In this section, the results of the calibration room's adjustment and the geometric calibration of the Panono camera (MPC_{36}) datasets 1 and 2 are presented.

The Calibration of SFCs and the Calibration Room's Geometry

A standard BBA was used to determine the geometry of the calibration room by employing the Canon EOS 6D and Samsung NX300 cameras, as well as for calibrating the two cameras. The self-calibrating BBA was performed by employing 10 parameters for the interior orientation of the SFCs: principal distance; principal point in the x and y directions; radial (K_1, K_2, K_3) and tangential (P_1, P_2) lens distortions; scale and shear.

The estimated sensor information, including a posteriori variances, for the Canon EOS 6D and Samsung NX300 cameras are listed in Table I. Most of the estimated distortion

values could be considered as significant when comparing the parameter values to their corresponding standard deviation values. Moreover, the significance of an individual distortion parameter was assessed by temporarily blocking it from the model to measure the largest displacement caused by it. The results of this test are demonstrated for the Canon EOS 6D and Samsung NX300 in Fig. 5. The scale factor in the standard BBA did not exhibit a significant effect on image residuals, which was also confirmed by the closeness of the standard deviation to the estimated value in Table I. The corresponding image displacement by removing this parameter (Fig. 5) also confirms the insignificance of this parameter; therefore, the SFCs were not noticeably affected by the conformity induced by this parameter. In the evaluation, all symmetric radial distortion parameters (K_1, K_2, K_3) were combined (as K_i) to avoid misleading the reader by large displacements caused by correlated parameters. The magnitude of the corresponding standard deviation values for scale and shear still suggested that the estimated values were meaningful, and thus all parameters were included in the final model. The results showed somewhat lower standard deviation values for the IOPs of the EOS 6D compared to those for the NX300.

TABLE I. The calibrated parameters for the Canon EOS 6D and Samsung NX300; *s.d.* = standard deviation.

			Canon EOS 6D		*Samsung NX300*	
No.	*Parameter*	*Units*	*Value*	*s.d.*	*Value*	*s.d.*
1	Principal distance	pixel	3157·65	1·22E−01	3806·53	2·76E−01
2	Principal distance	mm	20·65	7·96E−04	16·34	1·18E−03
3	Principal point x	pixel	2754·72	1·57E−01	2685·69	3·86E−01
4	Principal point y	pixel	1516·57	1·60E−01	1833·81	3·56E−01
5	K_1	–	8·34E−02	1·35E−04	1·57E−02	2·30E−04
6	K_2	–	−7·41E−02	3·57E−04	−2·53E−02	7·86E−04
7	K_3	–	1·06E−02	2·80E−04	1·26E−02	8·19E−04
8	P_1	–	3·66E−04	1·30E−05	−6·72E−04	2·50E−05
9	P_2	–	9·30E−05	1·50E−05	1·60E−03	2·80E−05
10	Scale factor	–	1·60E−05	1·10E−05	9·90E−05	1·80E−05
11	Shear factor	–	−1·06E−04	5·00E−06	−1·20E−05	8·00E−06
12	Pixel size	μm	6·542		2·63	

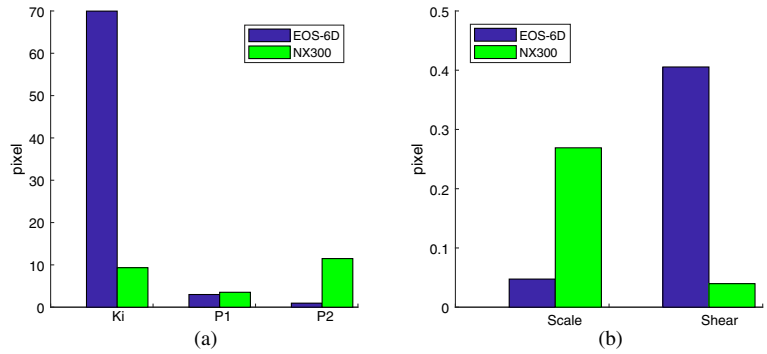


FIG. 5. Effect of blocking individual interior orientation parameters of the Canon EOS 6D and Samsung NX300 cameras. (a) Radial (K_i) and decentring (P_1, P_2) distortion. (b) Scale and shear.

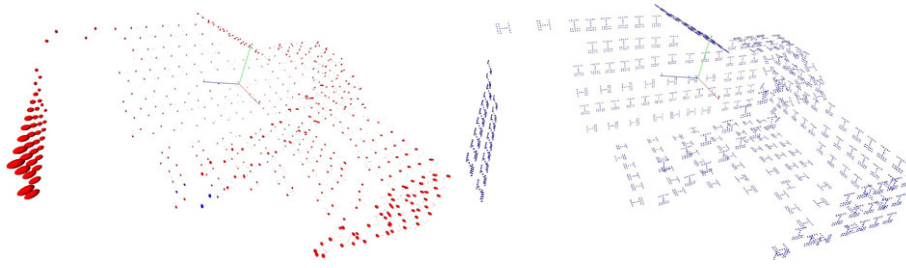
Important indicators of goodness of fit of the proposed model are the residuals of image observations; the average root mean square errors (RMSEs) in x and y directions are demonstrated in Fig. 8 for each image of the selected datasets. Most of the RMSE values for EOS 6D and NX300 were less than 0·3 pixel. The averages for EOS 6D dataset 6 were 0·15 and 0·14 pixel in the x and y directions, respectively: the corresponding figures for NX300 dataset 1 were 0·16 and 0·15, indicating slightly lower RMSEs for the EOS 6D.

In Fig. 6, the extracted structure of the calibration room and the 3D error ellipsoids of the corresponding 3D points are demonstrated. The estimated error ellipsoids shown in Fig. 6(a), demonstrate an acceptable fit of the BBA model to the observed data. As expected, stronger imaging geometry led to smaller error ellipsoids; for example, in Fig. 6(a) the points on the left have a lower number of intersecting rays, leading to larger error ellipsoids. A cumulative representation of the percentage of points with standard deviation values in predefined intervals is presented in Fig. 7. This histogram demonstrates the precision of the calibration room's structure, since most of the points (>90%) have standard deviation values less than 60 μm in all directions. The corresponding error ellipsoids of the Canon EOS 6D are smaller than those of the Samsung NX300; however, both cameras demonstrate acceptable qualities – such as sub-pixel image residuals, low RMSEs and small (<100 μm) error ellipsoids for the coded targets. The employment of error ellipsoids visibly helped in understanding weak intersections, low-accuracy cases and blunders; these played a fundamental role in the various implementations, including the standard and modified BBA. The RMSE of the differences between scale bars in EOS 6D dataset 7 and the ideal value was 0·2 mm, with a maximum difference of 0·6 mm.

By co-registering several datasets captured with both SFCs, acceptable registration residuals (RMSEs of 0·52, 0·51 and 0·67 mm in the X, Y and Z directions, respectively, with a confidence interval of 98%) were observed according to the standard Gaussian distribution that was assumed for the observations. The previous results presented precise geometric reconstruction of the calibration room with two SFCs.

MPC_{36} Calibration

Fraser (1982) stated that correlated parameters could weaken geometrically fragile networks and so threaten the convergence of Newton's iterative method. Linear correlations are addressed in a standard photogrammetric adjustment of SFC data by assuming a minimum set of constraints, and taking shots with varying scales (different positions and orientations). Variation in the depth of object points also helps to improve linear



(a) (b)

FIG. 6. The structure of the calibration room from the BBA. (a) Error ellipsoids, with the left-hand ellipsoids 300 times larger than those with better intersection geometry. (b) The location of points.

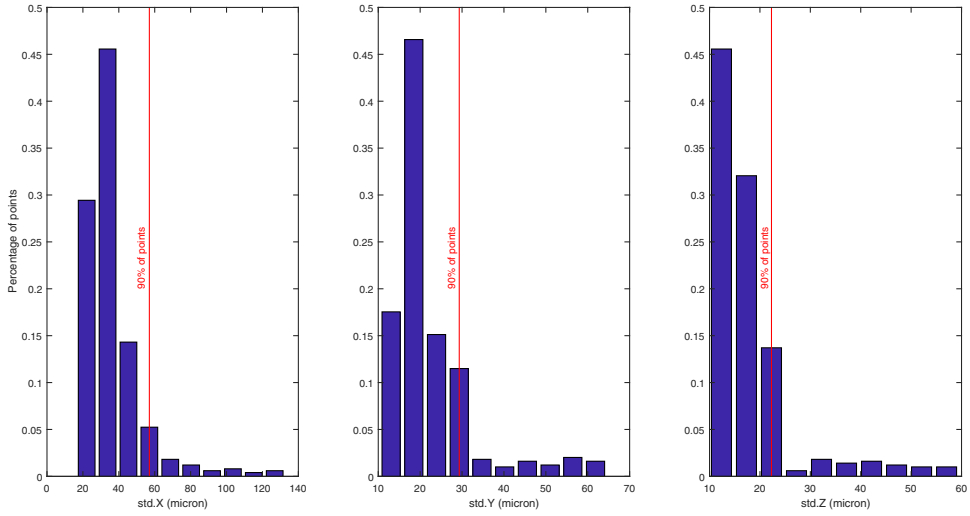


FIG. 7. Histograms of the targets' standard deviation values for dataset 7 of the Canon EOS 6D camera.

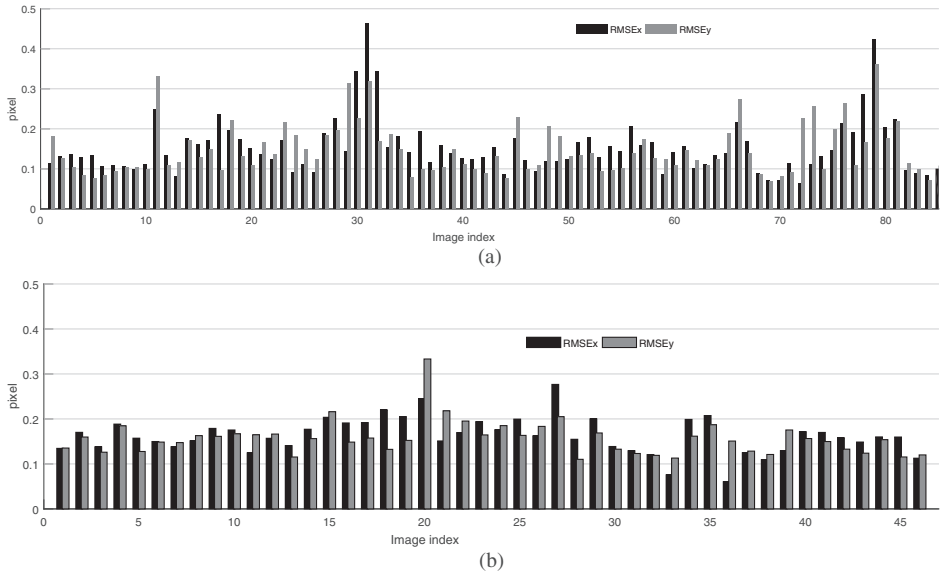


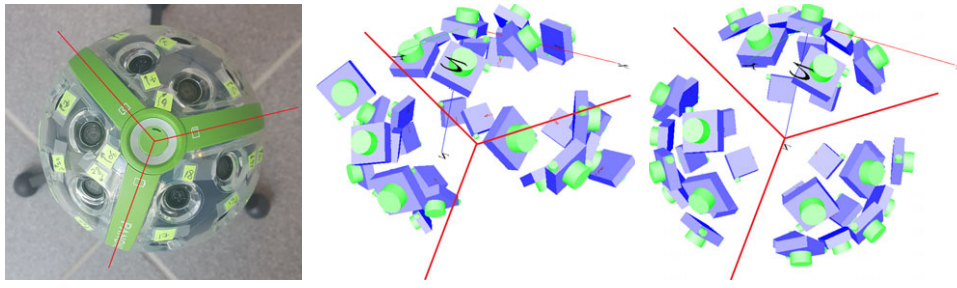
FIG. 8. RMSE of image coordinates in x (black) and y (grey) for: (a) Canon EOS 6D dataset 6; (b) Samsung NX300 dataset 1.

correlations. For an MPC with many side-looking cameras, however, the linearly correlated equations in the solution lead to a singularity that affects the convergence of the algorithm. If all the parameters of the cost function of an MPC with non-overlapping cameras are

considered free, **A** and **B** in equation (18) are singular specifically for a setting similar to that in this experiment where many cameras with opposing directions are attached to an MPC body. Other well-studied solutions to deal with this singularity, such as taking exposures with varying scale or depth variation, did not help in the case studied here. To overcome this problem, the selected solution was to reduce the number of degrees of freedom by considering the object points to be fixed and known, which was done by employing the calibration body. Thereby, the number of free parameters in equation (18) is decreased, which significantly improves the convergence of the algorithm. As a result, the modified BBA converged to a geometrically stable solution.

Initially, the primary structure of the underlying MPC_n was estimated through geometric averaging of a combined solution (equation (22); Fig. 9(b)). To enable this, a separate run of the standard BBA was performed on the individual projective cameras of the Panono. The RMSE values after running a standard BBA was of order 0·3 of a pixel. Since the calibrated structure that was recovered by the EOS 6D and NX300 proved to be of high quality and robust, the standard BBA was performed by considering the calibration object to be fixed and known. All 36 sets of camera location and orientations were combined into an initial MPC structure (Fig. 9(b)). This initial geometry deviated from reality (Fig. 9(a)) and was very noisy; the initial estimation of the sensor pose highlights the inability of the standard BBA to structurally estimate the underlying MPC_{36} (Fig. 9(b)). After running the modified BBA with the proposed sensor model (equations (15) to (21)), the final structure (Fig. 9(c)) resembled the reality of the sensor well. The average RMSEs of the image residuals of Panono datasets 1 and 2 in the x and y directions are given in Figs. 13(a) and (b), respectively. The magnitude of the Panono's RMSEs were slightly higher than those of the SFCs, with an average value of about 0·5 pixel. Both Panono datasets had somewhat comparable RMSE values. Fig. 10 demonstrates the pose (location and orientation) of Panono's shots in Panono datasets 1 and 2. The geometric constraints that were considered in both datasets, such as fixed height levels, as well as taking shots on straight lines in dataset 2, were used as a verification check. Obviously, the cluster of cameras in dataset 2 converged into linear clusters. In dataset 1, three different height levels were considered for the verification check, which later confirmed the correctness of the recovered structure.

For Panono, the significance test achieved by temporarily blocking a distortion parameter from the model is demonstrated in Fig. 12. The impact was greater for some cameras with poorer point distributions.



(a) (b) (c)

FIG. 9. Panono calibration. (a) Panono (top view). (b) The initial structure of the individual cameras. (c) The adjusted structure of the individual cameras.

Fig. 11 presents estimated principal distances of the Panono (for better visualisation the standard deviations are magnified 10 times). Fig. 14 shows the principal point locations with the corresponding 2D error ellipses. Variation in standard deviation values is mainly because of the inconsistency in the distribution of the coded targets in the calibration room. More specifically, larger error ellipses were observed for the upper and lower projective cameras, whereas the side-looking cameras' values were estimated more accurately. This inconsistency in estimated precisions is due to fewer coded targets on the ceiling and floor of the calibration room compared with the number of targets that were attached to the walls; it is aimed to improve the set-up of the targets in future work.

The calibrated ROPs, including a posteriori variances, for the Panono are listed in Table II, where lower accuracy cases are highlighted in grey. As expected, indices with a lower quality in their point distribution resulted in lower standard deviations ($i = 14, 19, 22, 23, 24$) compared to the indices with a good point coverage (middle cameras). For most of the indices, a standard deviation of less than 1′ was observed that indicates the

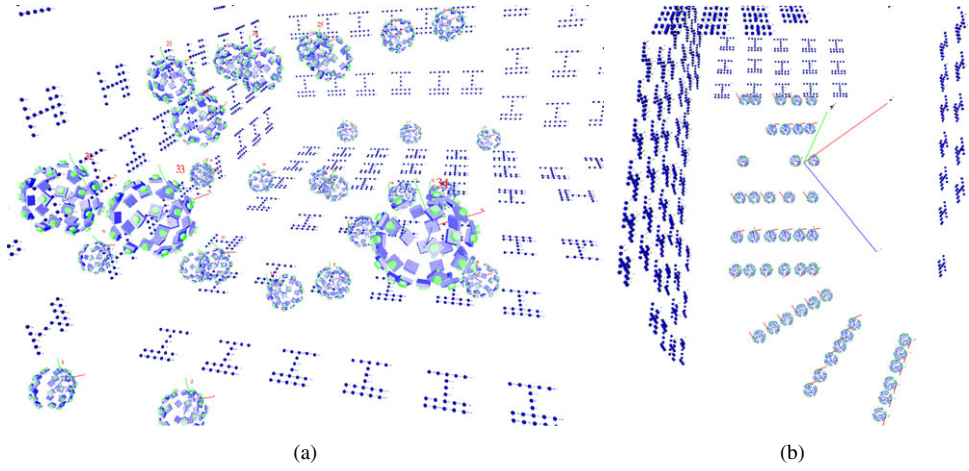


FIG. 10. Minimum constraint adjustment of the Panono camera (MPC_{36}): (a) dataset 1; (b) dataset 2.

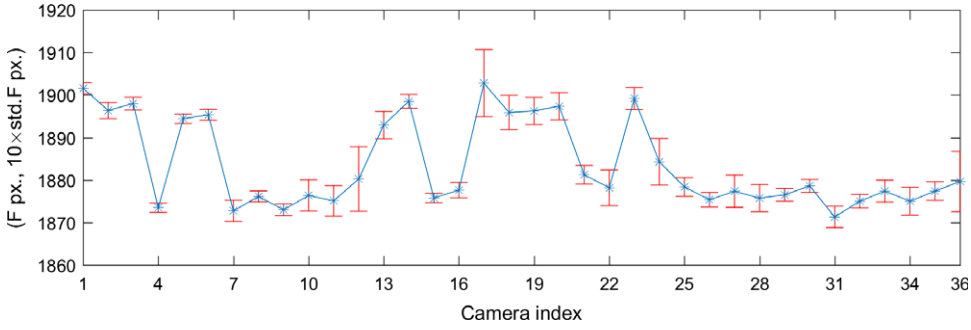


FIG. 11. Estimated values (blue) and standard deviations (red, with 10 times exaggeration) for the principal distances from Panono dataset 1.

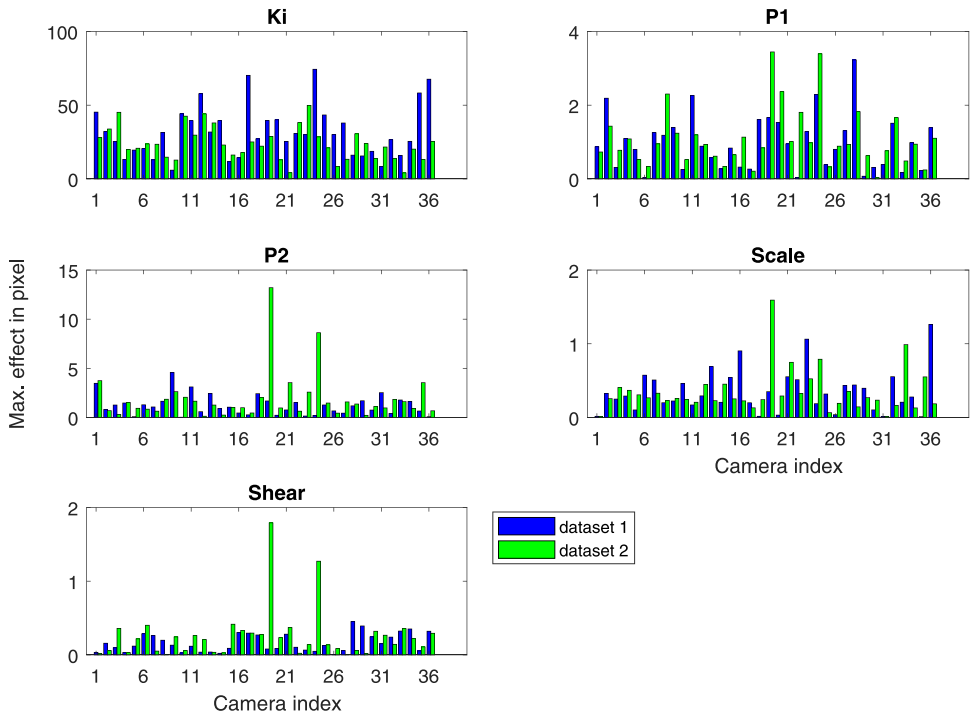


FIG. 12. Effect of blocking individual interior orientation parameters on Panono cameras.

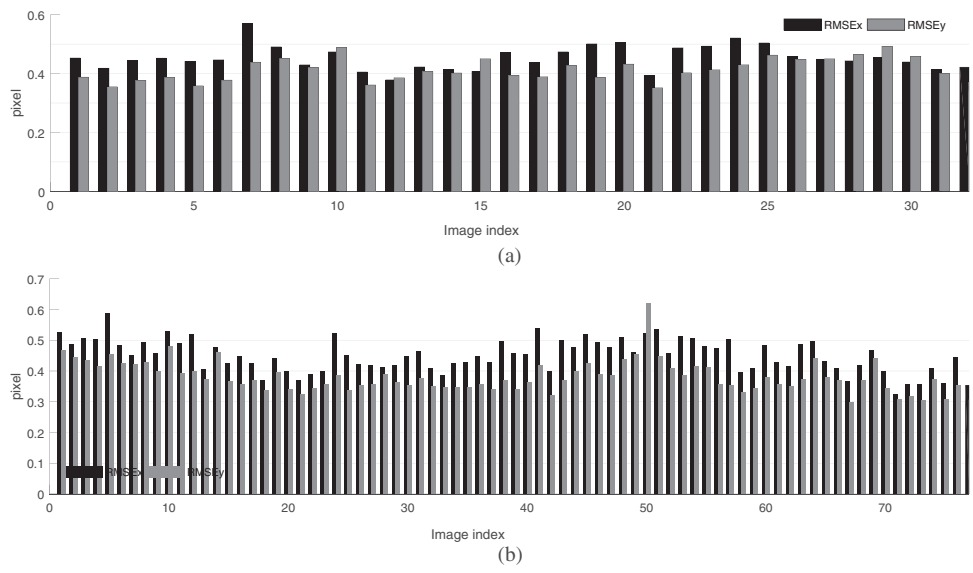


FIG. 13. RMSE of image coordinates (x and y) for: (a) Panono dataset 1; (b) Panono dataset 2.

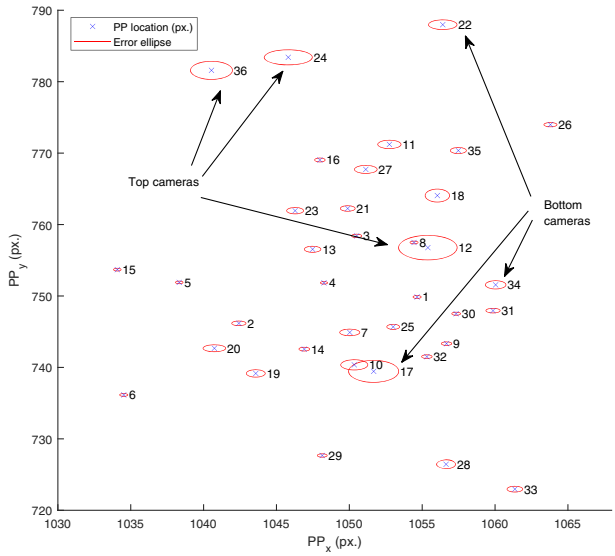


FIG. 14. Estimated values and error ellipses for the principal points of the Panono's individual projective sensors (dataset 1).

good quality of the adjustment. For relative positional accuracy, most of the cameras have a standard deviation less than 1 mm in all directions.

MPC_{36} Performance Assessment

The average RMSEs of image residuals in the x and y directions are given in Figs. 13(a) and (b) for Panono datasets 1 and 2, respectively. They are approximately two times larger than those of the Canon and Samsung. The standard deviations of the estimated object points derived from the Panono were less than 7·9 mm in all directions, compared to 0·5 mm for the EOS 6D and NX300. The differences between 3D check points and their corresponding intersected points from image points, the estimated calibration parameters and EOPs are presented in Fig. 15. In this figure, the differences are plotted as histograms in three main directions, and 90% error intervals are indicated by vertical red lines. In the X direction, the 90% error region corresponds to error between −10·60 and 11·23 mm, with an absolute mean of 5·41 mm and RMSE of 7·05 mm. In the Y direction, the 90% error zone is between −10·04 and 10·41 mm, with an absolute mean of 4·61 mm and RMSE of 6·17 mm. In the Z direction, the 90% error zone is between −12·39 and 13·31 mm, with an absolute mean of 5·91 mm and RMSE of 7·89 mm. The errors were close to the estimated standard deviation of the object points. The RMSE of the differences between scale bars in Panono dataset 2 and the reference value was 1·7 mm, with a maximum difference of 4·1 mm.

All these results were consistent and they indicated approximately 10 times poorer accuracy for the object point measurement by the Panono than those obtained by the good-quality Canon EOS 6D and Samsung NX300 SFCs. Poorer quality of the object reconstruction by the Panono could be seen in the image residuals and the standard deviations of estimated object points. Potential reasons for these poorer results for the

TABLE II. The calibrated relative orientation parameters for Panono projective cameras (i = 2:36). (ζ,η,ψ) are the relative Euler angles of the ith sensor with respect to the first sensor and (Δx, Δy, Δz) are the corresponding relative positions. Lower accuracy cases are highlighted in grey.

y	$\zeta°$	$\eta°$	$\psi°$	$\sigma(\zeta)'$	$\sigma(\eta)'$	$\sigma(\psi)'$	Δ_x cm	Δ_y cm	Δ_z cm	$\sigma(\Delta_x)$ mm	$\sigma(\Delta_y)$ mm	$\sigma(\Delta_z)$ mm
2	−33·16	22·65	6	0·51	0·63	0·28	−14·3595	−18·379	8·1692	0·788	0·312	0·254
3	−4·33	39·44	23·49	0·68	0·73	0·38	−29·508	−1·7369	9·1093	0·767	0·283	0·226
4	32·75	14·53	175·76	0·48	0·57	0·25	−9·2103	23·8215	5·7274	0·546	0·224	0·167
5	−2·67	−39·51	158·53	0·63	0·75	0·34	26·0584	0·1597	9·0407	0·604	0·253	0·187
6	38·03	−19·97	−88·34	0·64	0·49	0·32	14·5466	25·1199	8·9338	0·711	0·284	0·234
7	−60·65	5·05	−173·09	0·50	0·56	0·35	−5·7241	−38·902	23·3638	0·763	0·416	0·285
8	−28·24	−16·65	−175·52	0·53	0·61	0·22	7·9497	−24·2076	10·3893	0·634	0·23	0·19
9	−54·21	−34·68	−158·1	0·86	0·90	0·56	24·4724	−32·0512	24·0893	1·093	0·453	0·346
10	−92·61	−12·4	−1·66	0·80	0·93	0·44	4·6598	−46·2593	45·5489	1·03	0·485	0·304
11	−61·63	43·43	−19·68	0·84	0·80	0·69	−31·4242	−30·6216	29·5281	0·799	0·429	0·265
12	−100·07	23·06	94·33	0·89	0·55	0·55	−18·4274	−43·0956	50·3144	0·904	0·475	0·285
13	105·34	−62·59	157·09	1·20	0·47	1·18	45·1982	25·3175	61·8997	0·55	0·441	0·218
14	−143·58	−69·76	−98·44	5·35	1·04	5·15	47·8283	1·5496	67·6086	1·409	0·506	0·33
15	−26·68	−75·06	43·25	2·76	0·81	2·64	49·3087	−0·3107	45·1128	0·751	0·443	0·251
16	44·96	−49·06	−79·25	0·95	0·50	0·79	39·8615	29·7048	37·0432	0·84	0·416	0·27
17	123·13	−25·21	−30·02	0·50	0·61	0·41	26·3665	40·906	76·1823	0·579	0·389	0·223
18	80·75	−24·28	40·94	0·52	0·48	0·47	26·5201	45·8734	48·5735	0·724	0·46	0·268
19	−155·71	−36·96	63·42	8·09	8·55	4·95	34·9896	−11·4453	86·803	5·756	1·337	1·237
20	152·37	−43·17	23·16	0·72	0·81	0·54	36·4656	12·6709	84·4463	0·612	0·363	0·212
21	167·14	−16·18	59·44	1·20	0·96	0·33	18·8405	14·5956	99·9534	1·535	0·479	0·426
22	−156·74	−2·09	−118·28	1·61	1·76	0·20	11·7446	−17·2476	95·217	1·05	0·295	0·293
23	−104·63	−45·98	−103·39	3·27	1·73	2·36	39·3593	−27·3789	65·6346	1·234	0·503	0·324
24	−120·51	−12·09	−20·93	7·14	11·85	2·64	16·3223	−32·131	79·0035	6·996	0·991	0·92
25	100·22	35·29	−118·69	0·62	0·33	0·54	−33·6197	43·7468	59·162	0·47	0·404	0·195
26	146·78	39·17	−152·17	0·60	0·53	0·40	−35·1195	29·6184	79·3771	0·624	0·357	0·234
27	134·03	11·13	−121·15	0·53	0·46	0·31	−15·964	40·5492	85·087	0·642	0·351	0·232
28	95·25	0·32	54·61	0·51	0·38	0·40	−9·708	54·4846	55·5811	0·546	0·435	0·217
29	47·78	47·67	68·54	0·99	0·45	0·78	−39·7813	32·3196	34·7568	0·586	0·39	0·203
30	61·57	8·67	157·3	0·49	0·46	0·37	−17·183	48·1471	33·4619	0·53	0·349	0·181
31	−170·58	57·49	−14·83	1·08	0·85	0·88	−44·9157	3·9775	79·8133	1·015	0·454	0·331
32	94·86	67·91	71·13	1·79	0·44	1·67	−47·9375	20·5264	61·7758	0·578	0·44	0·212
33	−35·97	81·04	−129·57	5·71	0·87	5·68	−54·8891	1·9838	49·8137	1·43	0·604	0·441
34	−112·9	50·29	114·57	1·26	0·58	1·07	−41·4337	−20·0616	70·6331	1·035	0·505	0·343
35	176·19	19·12	152·09	0·77	0·80	0·26	−23·294	8·0405	97·2242	0·924	0·351	0·316
36	−143·33	21·16	−131·82	0·93	0·80	0·44	−25·8005	−17·8486	88·8922	1·09	0·437	0·36

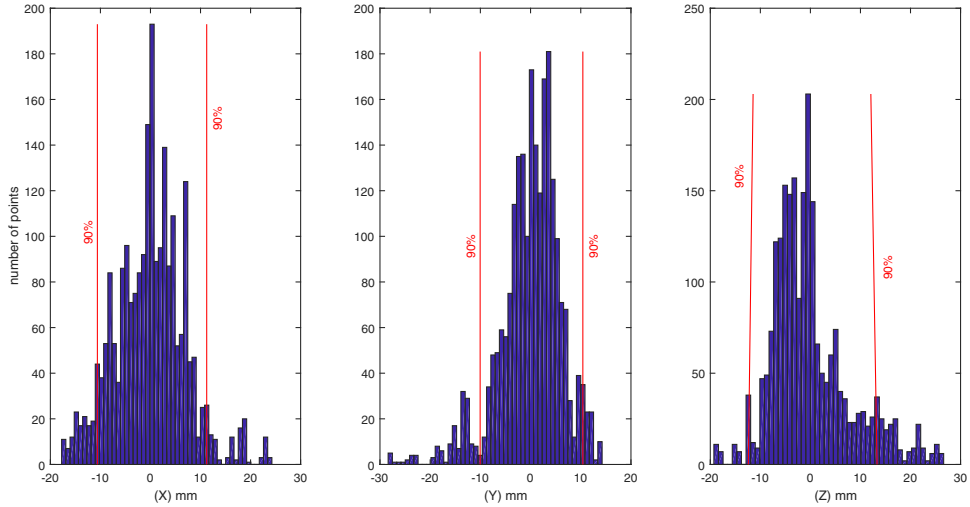


FIG. 15. Histograms of differences between ground truth from Canon EOS 6D dataset 7 and model points from Panono dataset 2. A total of 2300 3D check points were used after registration of model coordinate system to the ground in the *X*, *Y* and *Z* directions.

Panono include poorer image quality, a less than ideal FOV and poorer geometry of the image block. Furthermore, some inconsistencies remain in the sensor model and, in particular, some individual cameras having ROPs of poorer quality can cause a reduction in the overall accuracy. However, the results can be considered of appropriate quality, and the 360° coverage makes this inexpensive MPC an attractive target tool for many applications such as 3D reconstruction of indoor spaces and texturing.

SUMMARY

This paper has presented the proposed photogrammetric paradigm for the MPC calibration as a consistent model with the standard BBA. The algorithm has proved to be an efficient model for estimating the geometrical structure of a general MPC_n. The major contributions of this work include:

(1) Offering a uniform photogrammetric paradigm for $\text{MPC}_n (n \geq 2)$.
(2) Addressing the singularity in the calibration process of an $\text{MPC}_n (n \geq 2)$ where no overlap exists by proposing a fixed calibration body (a calibration room).
(3) Modifying the well-known BBA to accept a more general camera model as a solid single unit.
(4) Successfully demonstrating the real-case adjustment of a semi-panoramic MPC_{36} by employing the proposed model.

The efforts of the authors' future research will extend the algorithm to accept more general classes of cameras as well as further development of the calibration room to obtain accurate calibration of all cameras with a minimum number of images. It will also investigate the utilisation of various 360° cameras in photogrammetric reconstruction when used for mobile reconstruction tasks.

ACKNOWLEDGEMENTS

We are grateful to the anonymous reviewers whose comments helped us to improve the manuscript. This work was funded by the Academy of Finland through the "UAV 4D Bio" project (grant number 273806).

APPENDIX: SYMBOLS AND CONVENTIONS

Throughout this work, the following notation is used.

The Photogrammetric Record's normal conventions are maintained: bold for matrices ($\mathbf{A}$) and vectors ($\mathbf{x}$); $\mathbf{A}^{\mathrm{T}}$ and $\mathbf{A}^{-1}$ for the transpose and inverse of the positive-definite matrix $\mathbf{A}$; regular characters for scalars such as constants and variables; and $\mathbf{R}$ for a 3×3 rotation matrix.

In addition:

$(\mathbf{A})^{+}$ is a right pseudo inverse of the semi-positive definite matrix ($\mathbf{A}$), and ${}^{+}(\mathbf{A})$ is the left pseudo inverse.

A sub-matrix access is represented by $\mathbf{A}_{(r1:r2,c1:c2)}$ where r and c are rows and columns; the sub-index could be a single number if $\mathbf{A}$ is a one-dimensional vector.

A time index is a lower case subscript t such as $(\mathbf{X_0})_t$.

Variables in a range are in the form (variable = start index: end index) such as ($i = 1{:}10$).

REFERENCES

AMINI, A. S., VARSHOSAZ, M. and SAADATSERESHT, M., 2014. Development of a new stereo-panorama system based on off-the-shelf stereo cameras. *Photogrammetric Record*, 29(146): 206–223.

BOUGUET, J. Y., 2011. *Stereo Camera Calibration Toolbox*. https://www.vision.caltech.edu/bouguetj/calib_doc/htmls/example5.html [Accessed: 9th November 2017].

BROWN, D. C., 1966. Decentring distortion of lenses. *Photogrammetric Engineering*, 32(3): 444–462.

CLARKE, T. A. and FRYER, J. G., 1998. The development of camera calibration methods and models. *Photogrammetric Record*, 16(91): 51–66.

DESOUZA, G. N. and KAK, A. C., 2002. Vision for mobile robot navigation: a survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 24(2): 237–267.

ENGLISH, A., ROSS, P., BALL, D. and CORKE, P., 2014. Vision based guidance for robot navigation in agriculture. *IEEE International Conference on Robotics and Automation (ICRA)*, Hong Kong. Pages 1693–1698.

FRASER, C. S., 1982. On the use of non-metric cameras in analytical close range photogrammetry. *Canadian Surveyor*, 36(3): 259–279.

FRASER, C. S., 1997. Digital camera self-calibration. *ISPRS Journal of Photogrammetry and Remote Sensing*, 52(4): 149–159.

FRASER, C. S., 2013. Automatic camera calibration in close range photogrammetry. *Photogrammetric Engineering & Remote Sensing*, 79(4): 381–388.

FURUKAWA, Y. and PONCE, J., 2007. Accurate, dense, and robust multi-view stereopsis. *IEEE Conference on Computer Vision and Pattern Recognition*, Minneapolis, Minnesota, USA. 8 pages.

FURUKAWA, Y., CURLESS, B., SEITZ, S. M. and SZELISKI, R., 2010. Towards Internet-scale multi-view stereo. *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, San Francisco, California, USA. Pages 1434–1441.

GARRIDO-JURADO, S., MUÑOZ-SALINAS, R., MADRID-CUEVAS, F. J. and MARÍN-JIMÉNEZ, M. J., 2014. Automatic generation and detection of highly reliable fiducial markers under occlusion. *Pattern Recognition*, 47(6): 2280–2292.

GASPARINI, S. and BERTOLINO, P., 2013. Stereo camera tracking for mobile devices. *IEEE Conference on Computer Vision and Pattern Recognition Workshops*, Portland, Oregon, USA. Pages 14–19.

GRANSHAW, S. I., 1980. Bundle adjustment methods in engineering photogrammetry. *Photogrammetric Record*, 10(56): 181–207.

GROSSO, E. and TISTARELLI, M., 1993. Active/dynamic stereo: a general framework. *Proceedings of IEEE Conference on Computer Vision and Pattern Recognition*, New York, USA. Pages 732–734.

GRUEN, A. and HUANG, T. S. (Eds.), 2001. *Calibration and Orientation of Cameras in Computer Vision*. Springer, Berlin, Germany. 236 pages.

HABIB, A., DETCHEV, I. and KWAK, E., 2014. Stability analysis for a multi-camera photogrammetric system. *Sensors*, 14(8): 15084–15112.

HARTLEY, R. I., 1997. In defense of the eight-point algorithm. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 19(6): 580–593.

HARTLEY, R. I. and ZISSERMAN, A., 2004. *Multiple View Geometry in Computer Vision*. Second edition. Cambridge University Press, Cambridge, UK. 672 pages.

HE, G., NOVAK, K. and FENG, W., 1993. Stereo camera system calibration with relative orientation constraints. *SPIE*, 1820: https://doi.org/10.1117/12.1241367.

KHORAMSHAHI, E., HONKAVAARA, E. and ROSNELL, T., 2017. An automatic method for adjustment of a camera calibration room. *International Federation of Surveyors Working Week*, Helsinki, Finland. 9 pages.

LERMA, J. L., NAVARRO, S., CABRELLES, M. and SEGUÍ, A. E., 2010. Camera calibration with baseline distance constraints. *Photogrammetric Record*, 25(130): 140–158.

LI, B., HENG, L., KOSER, K. and POLLEFEYS, M., 2013. A multiple-camera system calibration toolbox using a feature descriptor-based calibration pattern. *IEEE/RSJ International Conference on Intelligent Robots and Systems*, Tokyo, Japan. Pages 1301–1307.

LICHTI, D. D., SHARMA, G. B., KUNTZE, G., MUND, B., BEVERIDGE, J. E. and RONSKY, J. L., 2015. Rigorous geometric self-calibrating bundle adjustment for a dual fluoroscopic imaging system. *IEEE Transactions on Medical Imaging*, 34(2): 589–598.

PARACCHINI, M., SCHEPIS, A., MARCON, M., FALCHETTO, M., PLEBANI, E. and PAU, D., 2016. Accurate omnidirectional multi-camera embedded structure from motion. *IEEE 2nd International Forum on Research and Technologies for Society and Industry*, Bologna, Italy. Pages 78–83.

PFINGSTHORN, M., RATHNAM, R., LUCZYNSKI, T. and BIRK, A., 2016. Full 3D navigation correction using low frequency visual tracking with a stereo camera. *IEEE Oceans Conference*, Shanghai, China. 6 pages.

SCHLEICHER, D., BERGASA, L. M., OCAÑA, M., BAREA, R. and LÓPEZ, M. E., 2009. Real-time hierarchical outdoor SLAM based on stereovision and GPS fusion. *IEEE Transactions on Intelligent Transportation Systems*, 10(3): 440–452.

SE, S., JASIOBEDZKI, P. and WILDES, R., 2007. Stereo-vision based 3D modeling of space structures. *SPIE 6555 (Sensors and Systems for Space Applications)*: 65550E1–65550E13.

SNAVELY, N., SEITZ, S. M. and SZELISKI, R., 2008. Modeling the world from Internet photo collections. *International Journal of Computer Vision*, 80(2): 189–210.

SVOBODA, T., HUG, H. and VAN GOOL, L., 2002. ViRoom – low cost synchronized multicamera system and its self-calibration. *Pattern Recognition: Lecture Notes in Computer Science*, 2449: 515–522.

SVOBODA, T., MARTINEC, D., PAJDLA, T., BOUGUET, J. Y., WERNER, T. and CHUM, O., 2011. *Multi-camera Self Calibration*. http://cmp.felk.cvut.cz/~svoboda/SelfCal/index.html [Accessed: 8th November 2017].

TOMMASELLI, A. M. G., GALO, M., DE MORAES, M. V. A., MARCATO JUNIOR, J., CALDEIRA, C. R. T. and LOPES, R. F., 2013. Generating virtual images from oblique frames. *Remote Sensing*, 5(4): 1875–1893.

TOMMASELLI, A. M. G., MARCATO JUNIOR, J., MORAES, M. V. A., SILVA, S. L. A. and ARTERO, A. O., 2014. Calibration of panoramic cameras with coded targets and a 3D calibration field. *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences*, 40(3/W1): 137–142.

TRIGGS, B., 1998. Autocalibration from planar scenes. *Lecture Notes in Computer Science*, 1406: 89–105.

TWEDDLE, B. E., 2010. *Computer Vision Based Navigation for Spacecraft Proximity Operations*. Master's thesis, Massachusetts Institute of Technology, Cambridge, Massachusetts, USA. 224 pages.

URBAN, S., WURSTHORN, S., LEITLOFF, J. and HINZ, S., 2017. MultiCol bundle adjustment: a generic method for pose estimation, simultaneous self-calibration and reconstruction for arbitrary multi-camera systems. *International Journal of Computer Vision*, 121(2): 234–252.

WANG, J., SUENAGA, H., HOSHI, K., YANG, L., KOBAYASHI, E., SAKUMA, I. and LIAO, H., 2014. Augmented reality navigation with automatic marker-free image registration using 3-D image overlay for dental surgery. *IEEE Transactions on Biomedical Engineering*, 61(4): 1295–1304.

WELLS, D. E. and KRAKIWSKY, E. J., 1971. *The Method of Least Squares*. Lecture Notes 18, Department of Geodesy and Geomatics Engineering, University of New Brunswick, Fredericton, Canada. 180 pages.

WU, C., 2013. Towards linear-time incremental structure from motion. *International Conference on 3D Vision*, Seattle, Washington, USA. Pages 127–134.

ZHANG, Z., 2000. A flexible new technique for camera calibration. *IEEE Transactions on Pattern Analysis and Machine Intelligence*, 22(11): 1330–1334.

ZHUANG, H., 1995. A self-calibration approach to extrinsic parameter estimation of stereo cameras. *Robotics and Autonomous Systems*, 15(3): 189–197.

Résumé

Les caméras multi-projectives (MPC), souvent basées sur le montage de plusieurs caméras sur un même support avec une faible base stéréoscopique et un recouvrement arbitraire, ont depuis peu fait leurs preuves dans les applications de la géomatique et de la vision par ordinateur. Cet article traite de l'étalonnage géométrique d'une MPC générique en présentant un modèle mathématique susceptible de décrire sa géométrie générique inconnue. Une compensation par blocs modifiée est utilisée pour étalonner une caméra 360° industrielle non métrique. La structure de n'importe quelle MPC peut être retrouvée à partir de l'ensemble des paramètres d'orientation relative et interne, ainsi que la position et l'orientation des caméras, au moyen d'une salle d'étalonnage préalablement déterminée avec précision par photogrammétrie rapprochée. Afin de démontrer l'efficacité et la précision du modèle, une caméra MPC Panono (constituée de 36 caméras individuelles) a été étalonnée. Au terme de l'ajustement, on observe des résidus sub-pixellaires dans l'espace image et des erreurs acceptables dans l'espace objet.

Zusammenfassung

In jüngster Zeit werden für Anwendungen in der Geomatik und Bildverarbeitung verstärkt Mehrfach-Projektive Kameras (MPCs) eingesetzt, die meist aus individuellen Kameras bestehen, die auf einer Plattform befestigt sind und eine kurze Basis und beliebige Überlappungen erlauben. Dieser Beitrag stellt Grundzüge einer geometrischen Kalibrierung einer solchen generellen Mehrfach-Projektiven Kamera in Form eines mathematischen Modells vor, das die unbekannte, generische Geometrie beschreibt. Eine angepasste Bündelblockausgleichung wird für die Kalibrierung einer nicht metrischen 360° Industriekamera eingesetzt. Die Struktur einer jeden MPC Kamera kann als ein Kalibriersatz von Parametern für Relative- und Innere Orientierung, bzw. auch für die Ausrichtung der einzelnen Aufnahmen, wiederhergestellt werden. Hierzu wird ein Kalibrierraum verwendet, der durch Nahbereichsphotogrammetrie hochgenau definiert worden ist. Zur Demonstration der Effizienz und Genauigkeit des angewandten Modells, wurde eine Panono Kamera (eine MPC mit 36 individuellen Kameras) kalibriert. Nach der Ausgleichung konnten Verbesserungen der Bildmessungen im Sub-Pixel Bereich und akzeptable Fehler im Objektraum festgetellt werden.

Resumen

Recientemente, las cámaras multi-proyectivas (MPCs), con frecuencia múltiples cámaras montadas en un soporte con línea de base pequeña y superposición arbitraria, han encontrado un espacio notable en la geomática y las aplicaciones de visión por ordenador. Este documento describe la calibración geométrica de un MPC general presentando un modelo matemático que describe su geometría genérica. Se emplea un ajuste de haces modificado para calibrar una cámara industrial no métrica de 360°. La estructura de cualquier MPC se puede recuperar como un conjunto de parámetros calibrados de orientación relativa e interior (así como la orientación de las captaciones MPC) usando una sala de calibración que se ha determinado con precisión mediante fotogrametría de objeto cercano. Para demostrar la eficiencia y precisión del modelo, se calibró una cámara Panono (un MPC con 36 cámaras individuales). Tras del ajuste, se observaron residuos subpíxel en las imágenes y errores aceptables en el espacio objeto.

摘要

近年来，采用短基线、任意重迭、基于框幅式组合的多投影相机组 (MPC) 在测绘及基于影像的测量应用中，有着显著的地位。本文描述一个通用的 MPC 几何校准方式，提出一个一般性未知几何的数学模型。本研究采用改进了的光束法区域网平差方法，来校准工业级别的 360°非量测性相机。任何 MPC 的组成结构以及镜头的姿态，都可以藉由近景摄影测量方式于校准室获得一组精确检校了的相对方位以及内部方位参数。为了展示该模型的效率以及精度，对一台 Panono 相机 (带有 36 个独立相机的 MPC) 进行了检校。平差后，获得子像元等级的像方残差，物方误差也达到可接受的标准。